A Sentiment Analysis Method to Better Utilize User Profile and Product Information

Capstone Project Presentation

Mingyu MA (Derek)

derek.ma@connect.polyu.hk

BSc (Hons) Computing, 14110562D

Supervisor: Prof. Qin LU

Co-examiner: Dr. Ajay Kumar PATHAK

2nd Assessor: Dr. Richard LUI

The Hong Kong Polytechnic University

香港理工大學

Department of Computing

電子計算學系

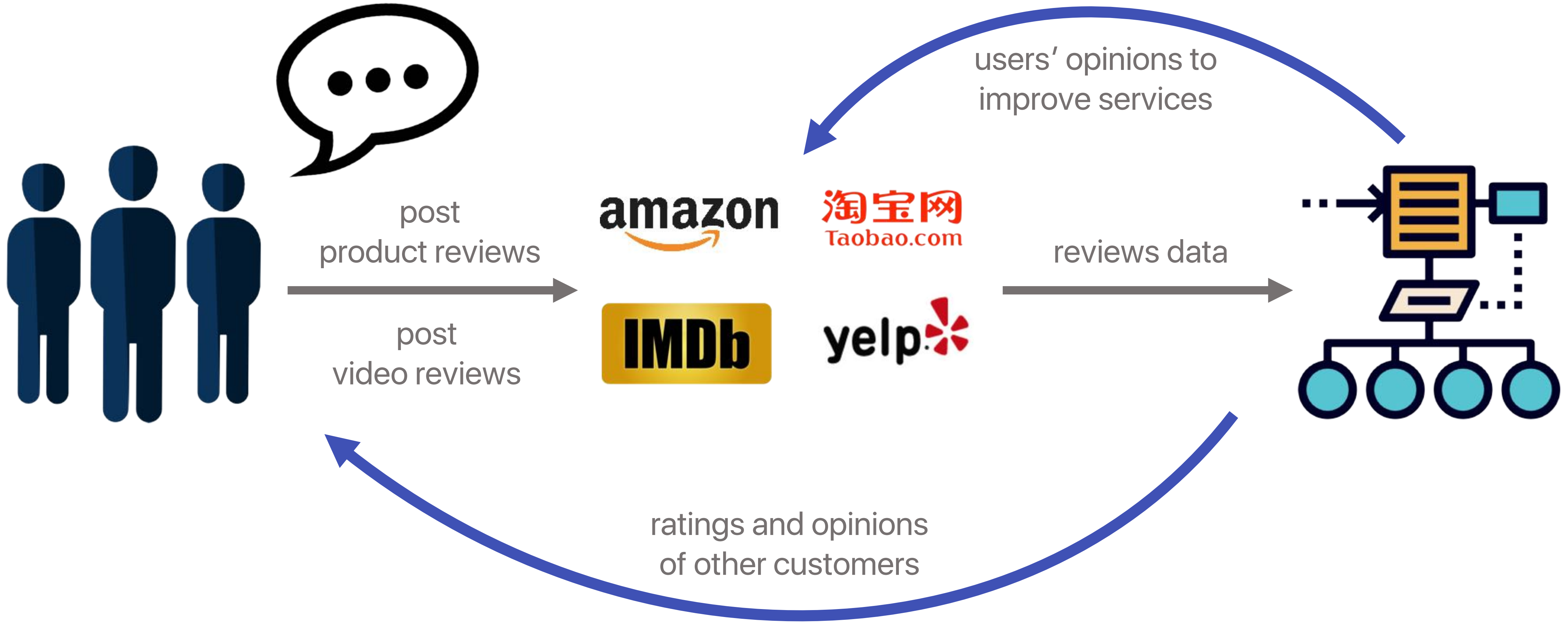
Contents

- Introduction
- Related Work
- Model Design
- Evaluation and Analysis
- Conclusion and Future Work

Introduction

Businesses would like to know users' opinions

Users can be benefited from others' opinions

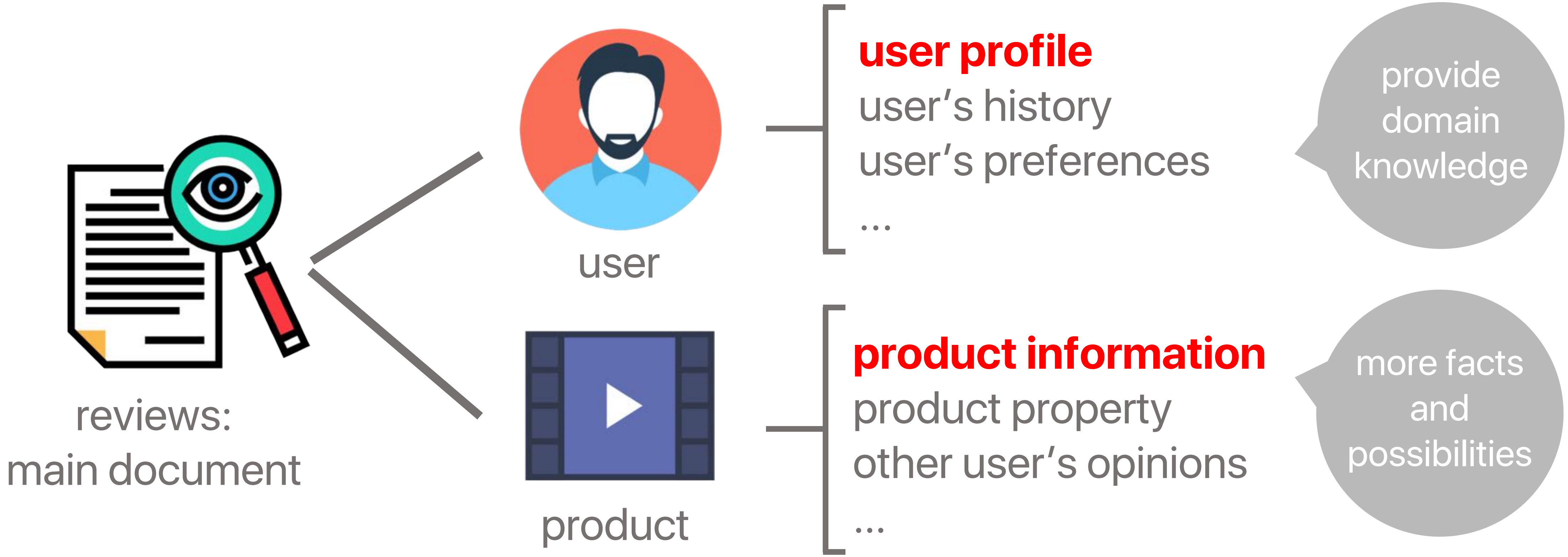


Introduction

Sentiment Analysis

methods of detecting, analyzing, and evaluating people's state of mind towards events, issues, or any other interest. (Yadollahi et al., 2017)

Background Info Is Available



Background Information Is Not Unified

- User’s perspective
  - Mean/lenient user
- Product’s perspective
  - Type, category
- **Different** background information influences the results in **different** perspectives

Objectives

A new sentiment analysis model

- **utilize** user and product information
- reflect impacts from user profile and product information **separately**

Related Work

Machine-Learning-based Sentiment Analysis

(Wang and Manning, 2012)

Linear model or kernel methods on lexical features

Traditional Way

(Tang et al., 2014), (Kim, 2014)

NN as classifier for text classification

RNN, LSTM

Neural-network-based Approaches

(Yang et al., 2016)
(Long et al., 2017)

Focus more on important text and add more associate data like eye-tracking data

Attention

User and Product Info in Sentiment Analysis

Utilizing **User Profile** and **Product Information** in Sentiment Analysis

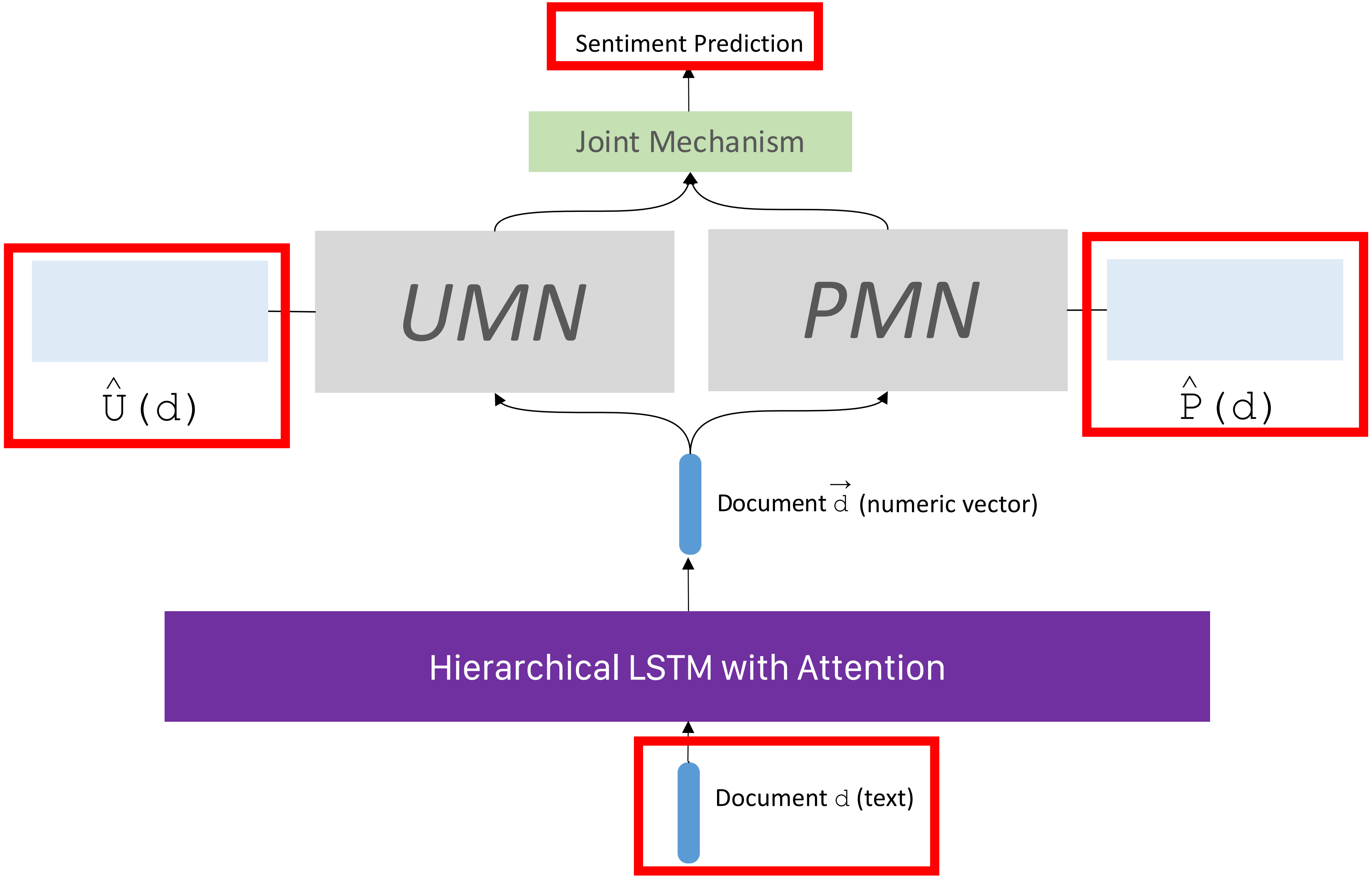
- **Memory network**
  (Tang, Qin and Liu, 2015; Dou, 2017)
  - RNN + external memory
- **Use external info as attention**
  (Chen et al., 2016)
  - State-of-the-art
- All consider user profile and product information as **single representation**

Model Design

JUPMN

Joint User and Product Memory Network

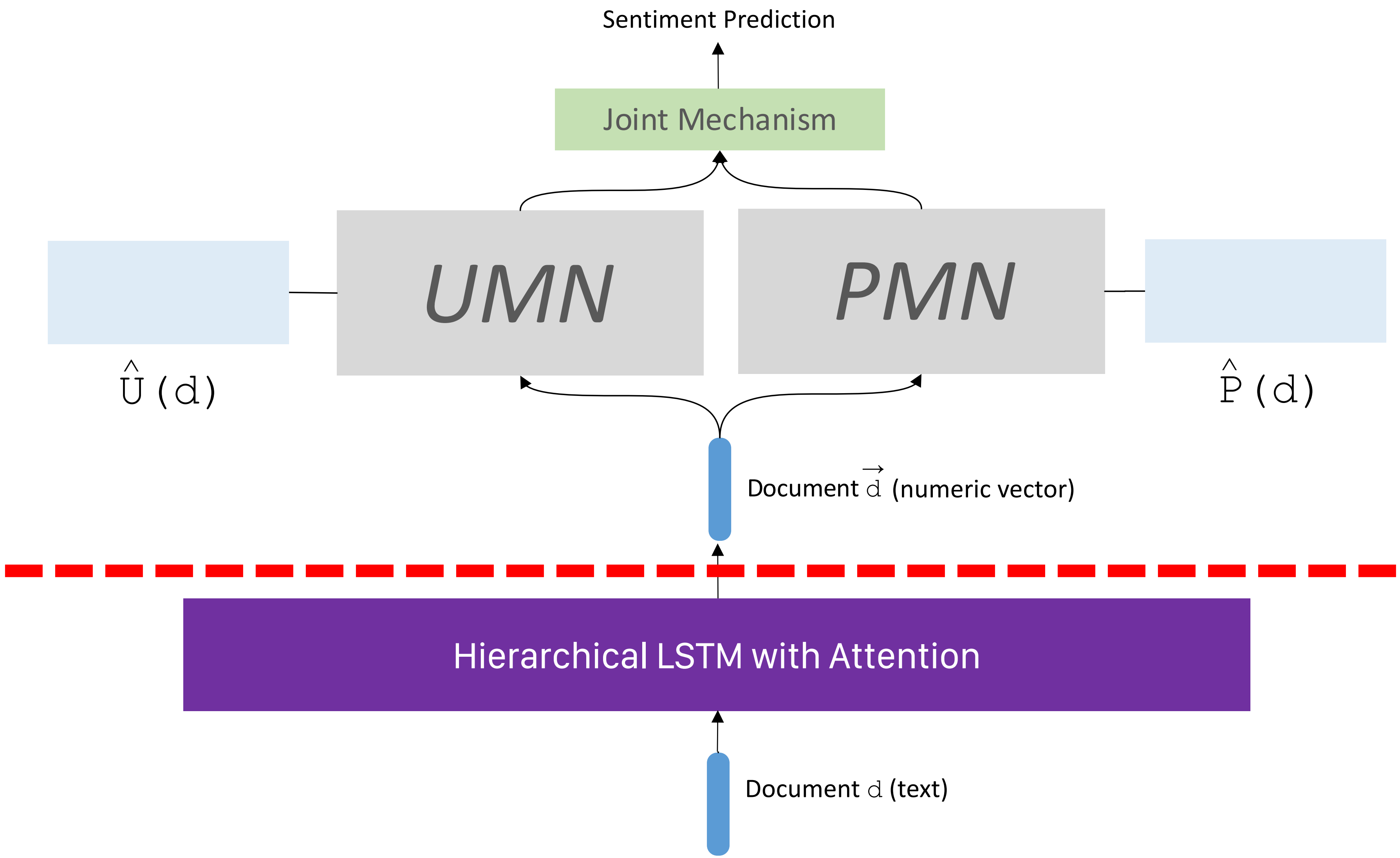
Model Overview



Input & Output

- Input
  - Document **d**
  - A writer **u**
  - A target **p**
- Output
  - Discrete sentiment prediction

Model Overview

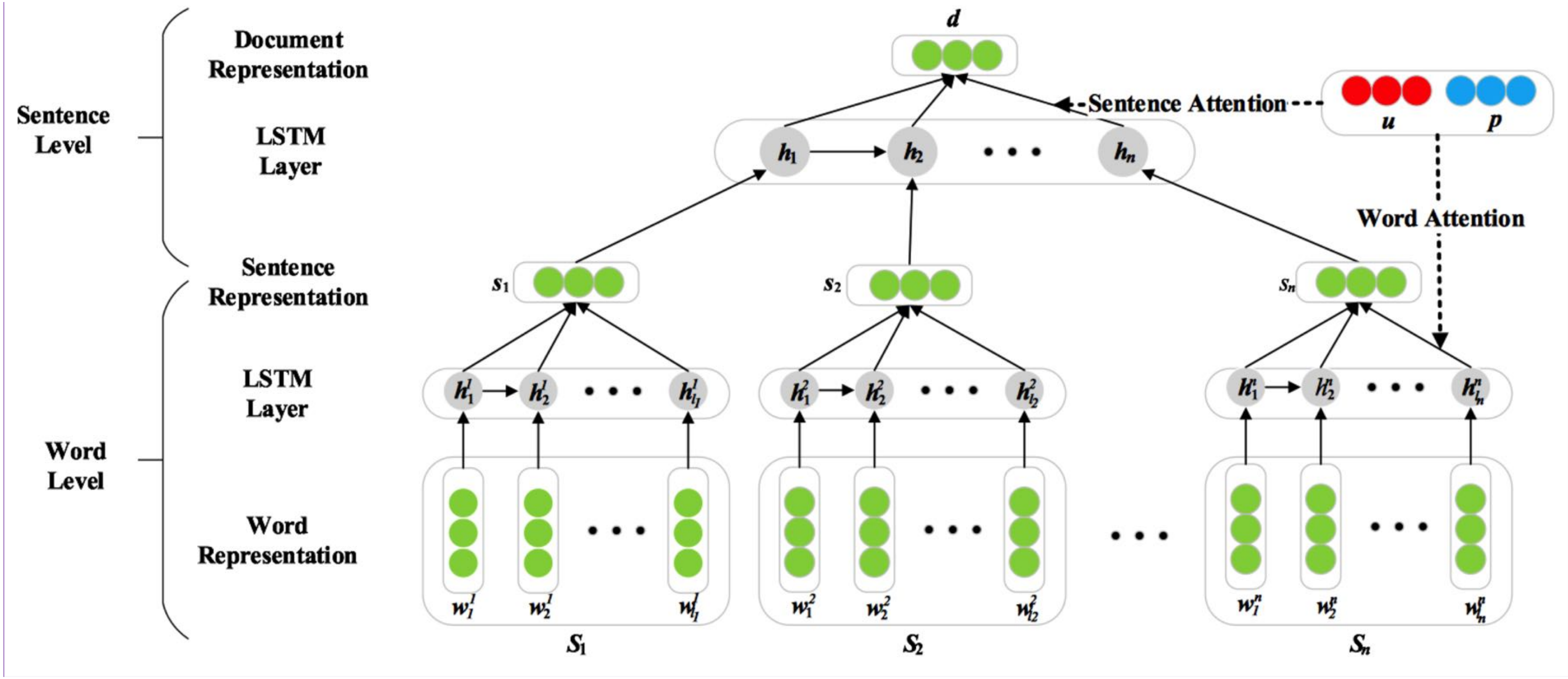


Structure

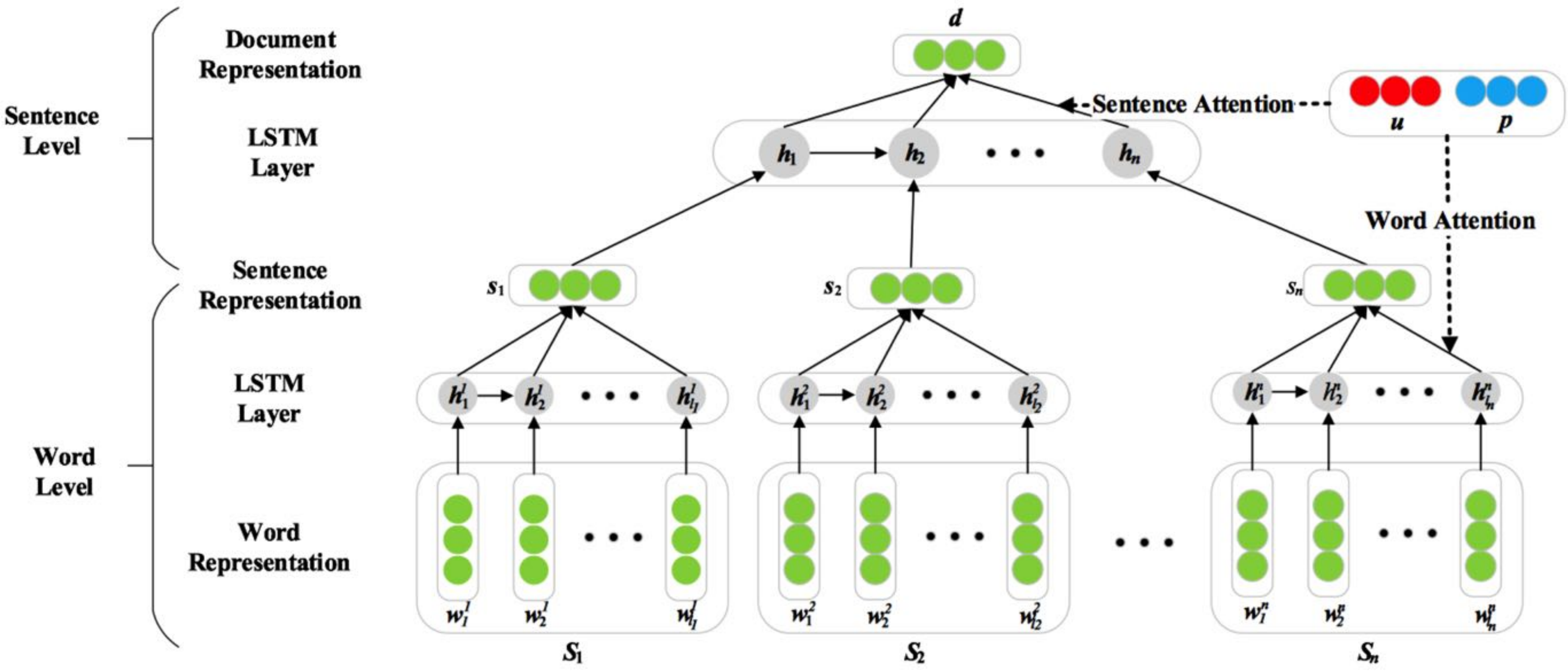
Part 2: Memory Networks

Part 1: Document Embedding

Hierarchical Long Short-Term Memory Network



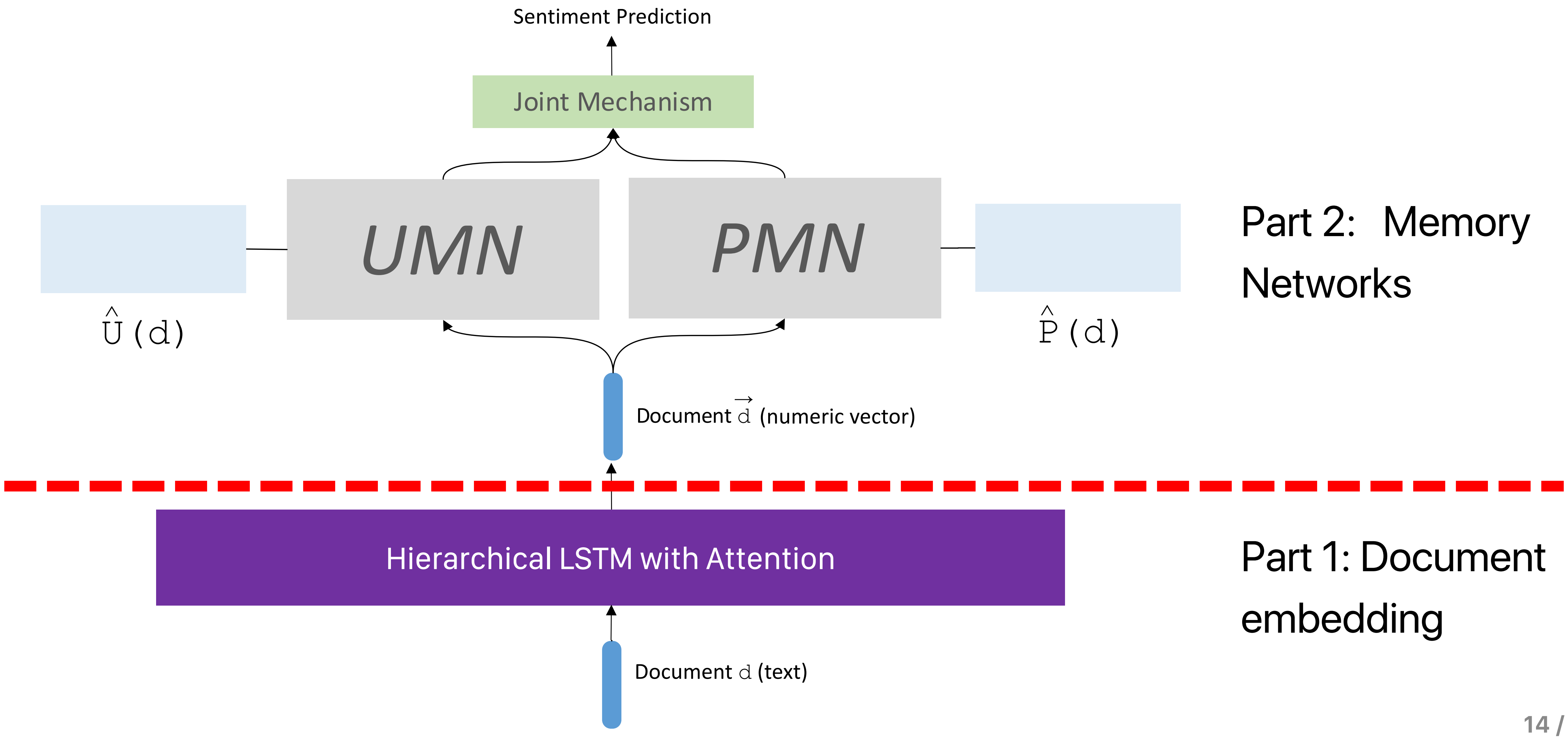
Hierarchical Long Short-Term Memory Network

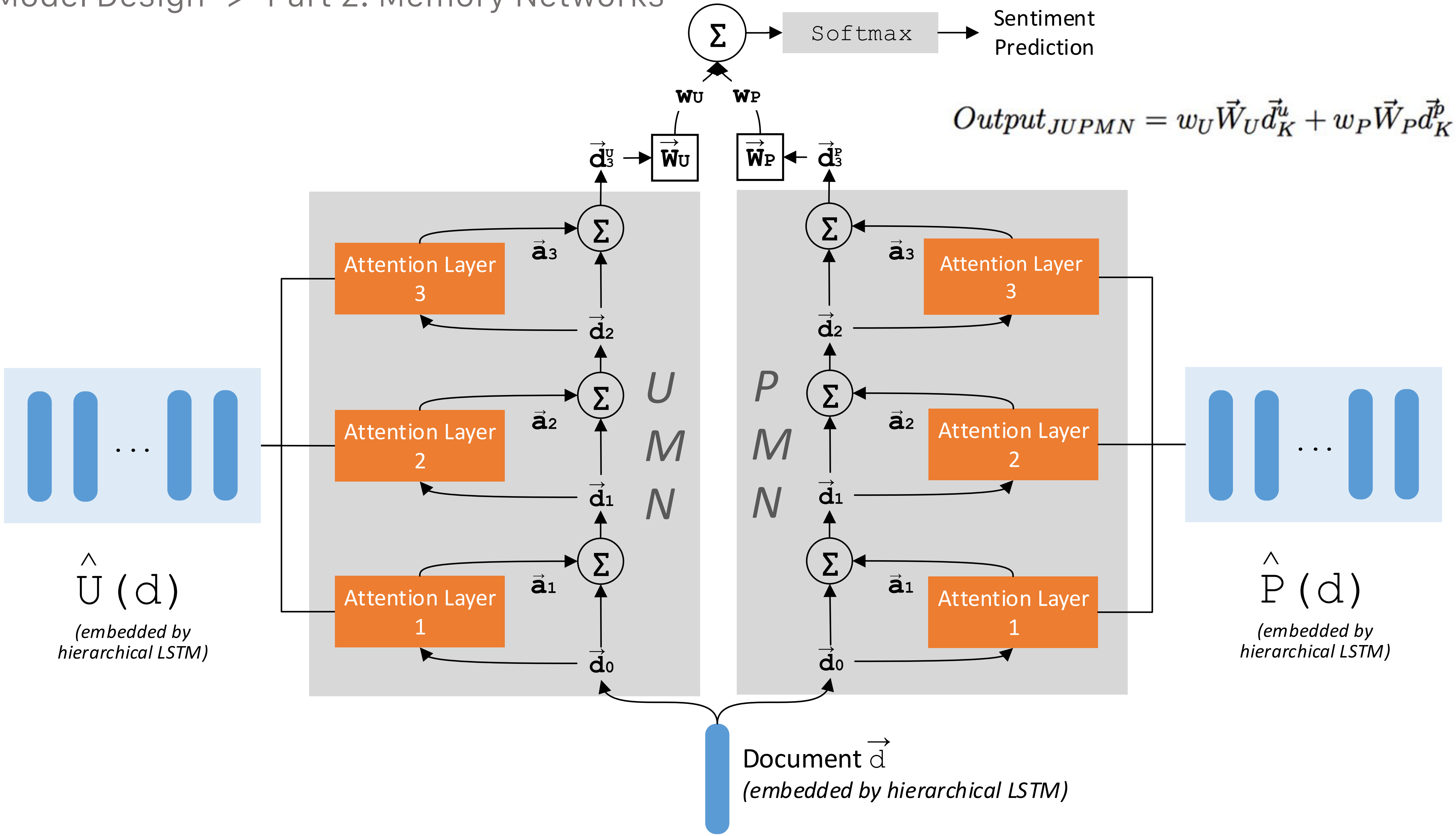


- Word-sentence-document level convention (Chen et al., 2016)
- Add attention in LSTM layers
  - With user and product attention
  - With eye-tracking cognition attention

Model Design

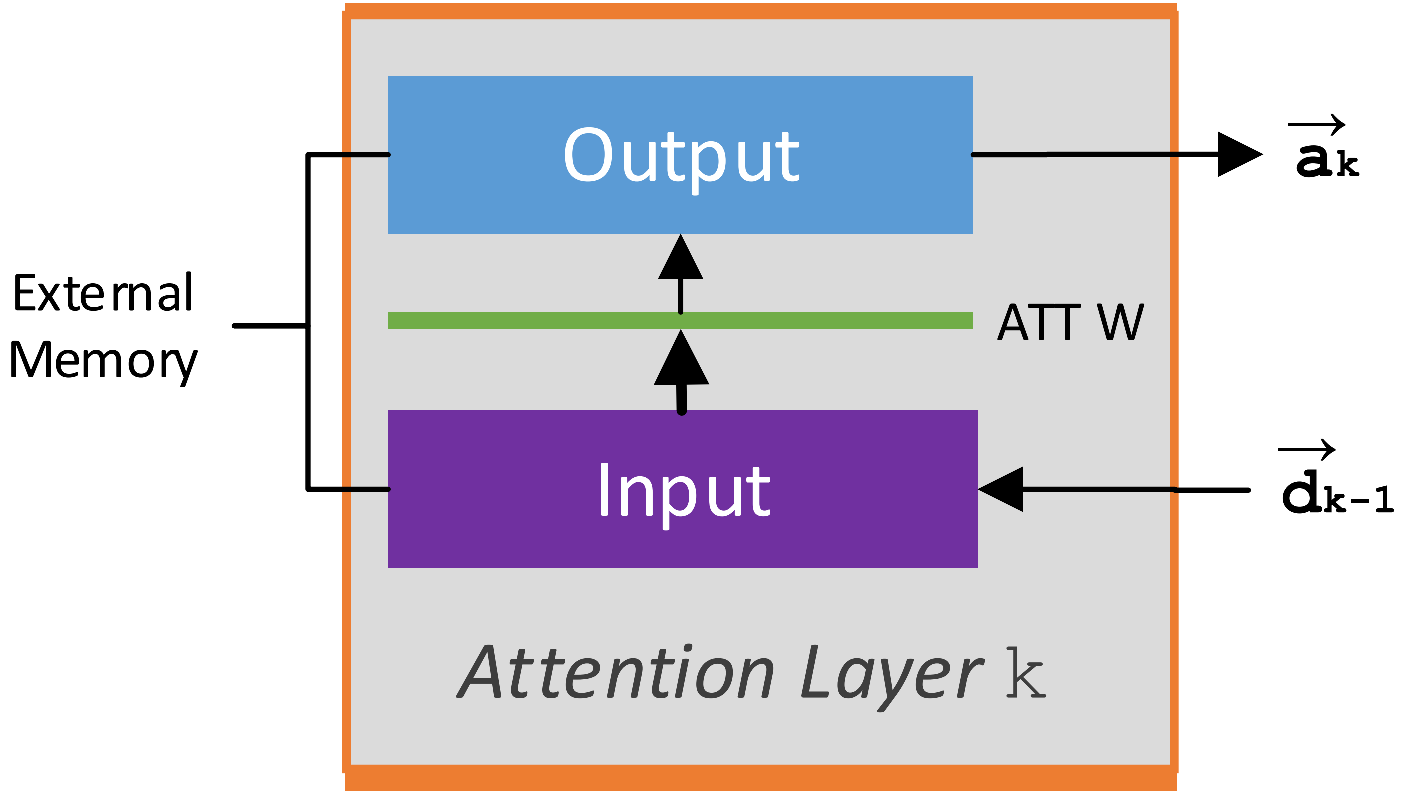
Part 2: Memory Networks





$$Output_{JUPMN} = w_U \vec{W}_U \vec{d}_K^u + w_P \vec{W}_P \vec{d}_K^p$$

Structure of Attention Layers



- Attention weight

$$\vec{p}_k = Softmax(\vec{d}_{k-1}^T * \hat{M})$$

- Output of attention layer

$$\vec{a}_k = \sum_{i=0}^{m} p_{ki} * \vec{M}_i.$$

Benchmark Datasets and Performance Metrics

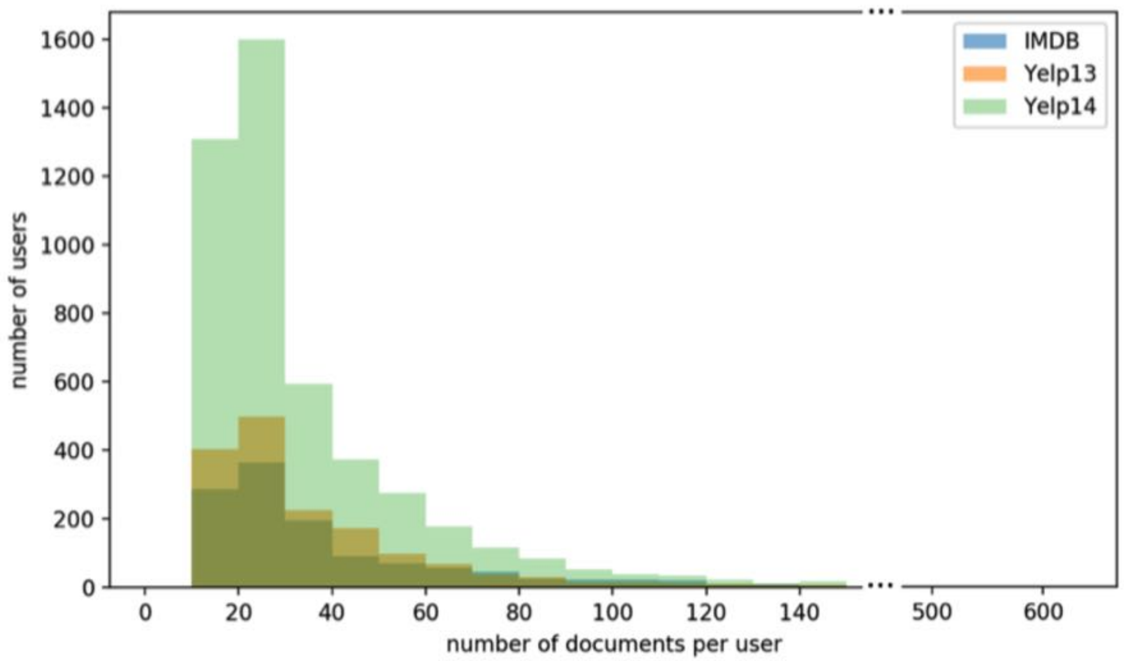
Three Benchmark Datasets

- IMDB
  - Diao et al., 2014
- Yelp 13, Yelp 14
  - Tang et al., 2015a

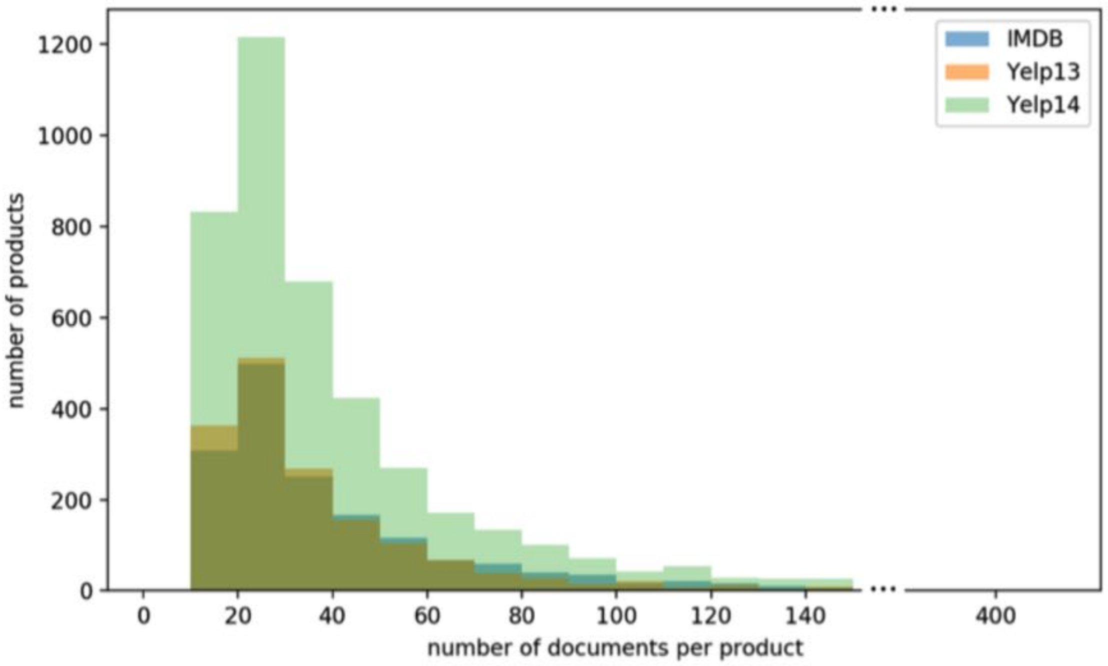
	IMDB	Yelp13	Yelp14
number of classes	10	5	5
number of review documents	84,919	78,966	231,163
number of users	1,310	1,631	4,818
number of products	1,635	1,631	4,194
average sentences' length	24.56	17.37	17.25

Benchmark Datasets and Performance Metrics

Three Benchmark Datasets



(a) Statistic of documents # per user



(b) Statistic of documents # per product

Benchmark Datasets and Performance Metrics

Performance Metrics

$$Accuracy = \frac{T}{N}$$

$$MAE = \frac{\sum_i |py_i - gy_i|}{N}$$

$$RMSE = \sqrt{\frac{\sum_i (py_i - gy_i)^2}{N}}$$

JUPMN and Comparison Models

Experimental Results

	IMDB			Yelp13			Yelp14		
Model	Acc	RMSE	MAE	Acc	RMSE	MAE	Acc	RMSE	MAE
Majority	0.196	2.495	1.838	0.392	1.097	0.779	0.411	1.06	0.744
Trigram	0.399	1.783	1.147	0.577	0.804	0.487	0.569	0.814	0.513
TextFeature	0.402	1.793	1.134	0.572	0.800	0.490	0.556	0.845	0.520
AvgWordvec	0.304	1.985	1.361	0.530	0.893	0.562	0.526	0.898	0.568
SSWE	0.312	1.973	N/A	0.549	0.849	N/A	0.557	0.851	N/A
RNTN+RNN	0.400	1.734	N/A	0.574	0.804	N/A	0.582	0.821	N/A
CLSTM	0.421	1.549	N/A	0.592	0.729	N/A	0.637	0.686	N/A
LSTM+LA	0.443	1.465	N/A	0.627	0.701	N/A	0.637	0.686	N/A
LSTM+CBA	0.489	1.365	N/A	0.638	0.697	N/A	0.641	0.678	N/A
UPNN(K)	0.435	1.602	0.979	0.608	0.764	0.447	0.596	0.784	0.464
UPDMN(K)	0.465	1.351	0.853	0.613	0.720	0.425	0.639	0.662	0.369
InterSub	0.476	1.392	N/A	0.623	0.714	N/A	0.635	0.690	N/A
LSTM+UPA	0.533	**1.281**	N/A	0.650	0.692	N/A	0.667	0.654	N/A
JUPMN	**0.539**	1.283	**0.725**	**0.662**	**0.667**	**0.375**	**0.676**	**0.641**	**0.351**

Group 1: simple methods based on language features

Group 2: models using machine learning

Group 3: models with user profile and product information in machine learning

JUPMN and Comparison Models

Experimental Results

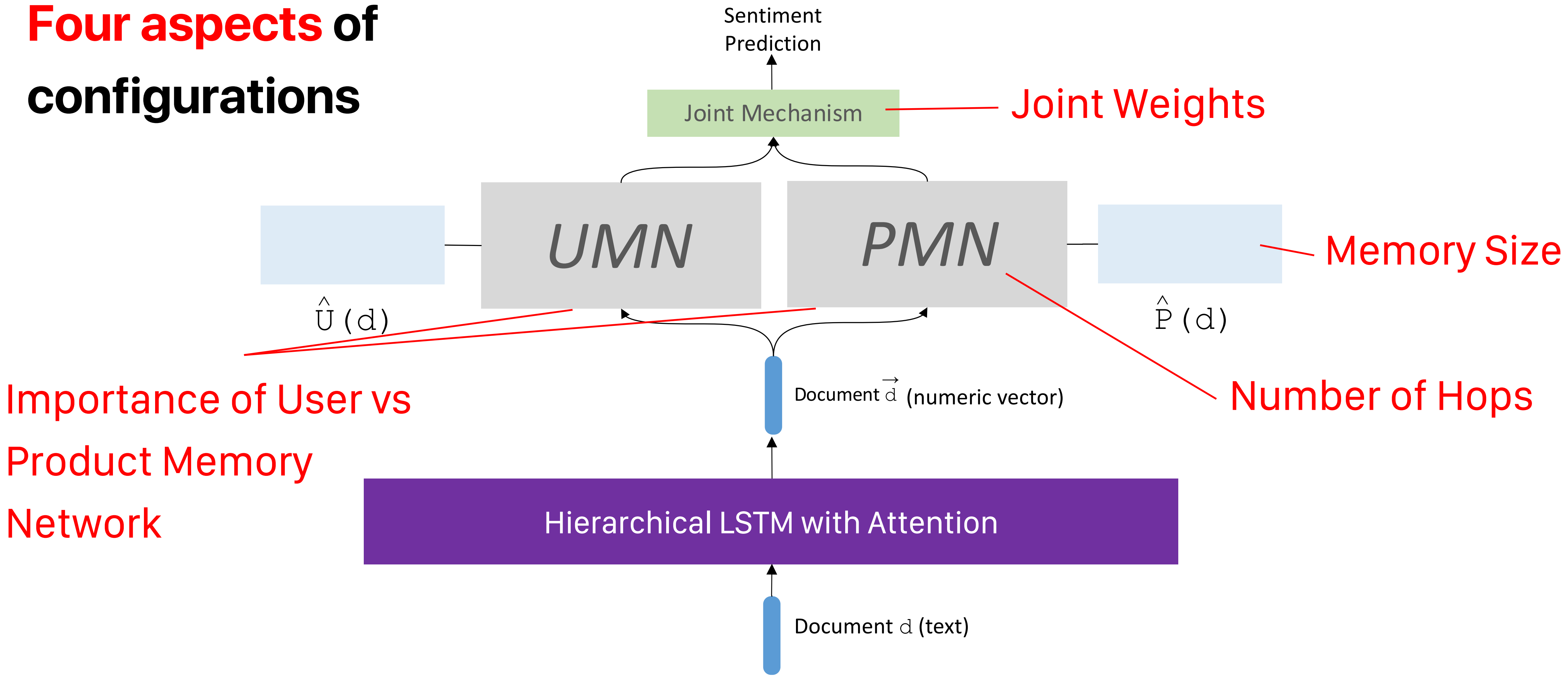
	IMDB			Yelp13			Yelp14		
Model	Acc	RMSE	MAE	Acc	RMSE	MAE	Acc	RMSE	MAE
Majority	0.196	2.495	1.838	0.392	1.097	0.779	0.411	1.06	0.744
Trigram	0.399	1.783	1.147	0.577	0.804	0.487	0.569	0.814	0.513
TextFeature	<u>0.402</u>	<u>1.793</u>	<u>1.134</u>	<u>0.572</u>	<u>0.800</u>	<u>0.490</u>	<u>0.556</u>	<u>0.845</u>	<u>0.520</u>
AvgWordvec	0.304	1.985	1.361	0.530	0.893	0.562	0.526	0.898	0.568
SSWE	0.312	1.973	N/A	0.549	0.849	N/A	0.557	0.851	N/A
RNTN+RNN	0.400	1.734	N/A	0.574	0.804	N/A	0.582	0.821	N/A
CLSTM	0.421	1.549	N/A	0.592	0.729	N/A	0.637	0.686	N/A
LSTM+LA	0.443	1.465	N/A	0.627	0.701	N/A	0.637	0.686	N/A
LSTM+CBA	<u>0.489</u>	<u>1.365</u>	N/A	<u>0.638</u>	<u>0.697</u>	N/A	<u>0.641</u>	<u>0.678</u>	N/A
UPNN(K)	0.435	1.602	0.979	0.608	0.764	0.447	0.596	0.784	0.464
UPDMN(K)	0.465	1.351	0.853	0.613	0.720	0.425	0.639	0.662	0.369
InterSub	0.476	1.392	N/A	0.623	0.714	N/A	0.635	0.690	N/A
LSTM+UPA	<u>0.533</u>	**1.281**	N/A	<u>0.650</u>	<u>0.692</u>	N/A	<u>0.667</u>	<u>0.654</u>	N/A
JUPMN	**0.539**	<u>1.283</u>	**0.725**	**0.662**	**0.667**	**0.375**	**0.676**	**0.641**	**0.351**

Findings

- JUPMN **outperforms** the state-of-the-art model
- Generally Group 2 performs better than Group 1, Group 3 performs better than Group 2
- Exceptions exist
  - *TextFeature*
  - *LSTM+CBA*

JUPMN with Different Configurations

Four aspects of configurations



Importance of User vs Product Memory Network

Experimental Results

	IMDB			Yelp13			Yelp14		
	Acc	RMSE	MAE	Acc	RMSE	MAE	Acc	RMSE	MAE
JUPMN-U(1)	0.536	**1.283**	0.737	0.656	0.687	0.380	0.667	0.655	0.361
JUPMN-U(2)	0.526	1.285	0.748	0.653	0.689	0.382	0.665	0.661	0.369
JUPMN-U(3)	0.524	1.295	0.754	0.651	0.692	0.388	0.661	0.667	0.374
JUPMN-P(1)	0.523	1.346	0.769	0.660	0.668	**0.370**	0.670	0.649	0.357
JUPMN-P(2)	0.517	1.348	0.775	0.656	0.680	0.380	0.667	0.656	0.364
JUPMN-P(3)	0.512	1.356	0.661	0.651	0.699	0.388	0.661	0.661	0.370
JUPMN(1)	**0.539**	**1.283**	**0.725**	**0.662**	**0.667**	0.375	**0.676**	**0.641**	**0.351**
JUPMN(2)	0.522	1.299	0.758	0.650	0.700	0.390	0.667	0.650	0.359
JUPMN(3)	0.502	1.431	0.830	0.653	0.686	0.382	0.658	0.668	0.371

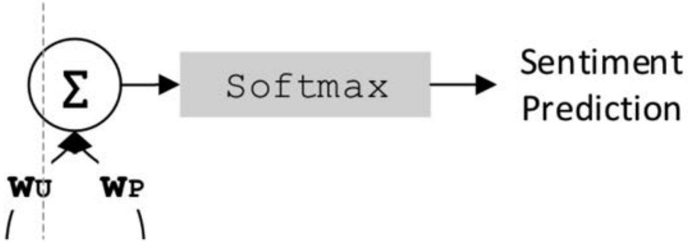
- JUPMN-U
  - With only User Memory Network
- JUPMN-P
  - With only Product Memory Network

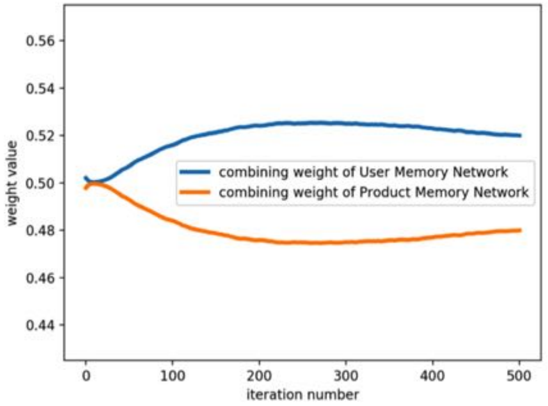
Observations

- User profile influences sentiments of movie reviews more
- Product information influences sentiments of restaurants reviews more

Importance of User vs Product Memory Network

Investigating by Checking Joint Weights

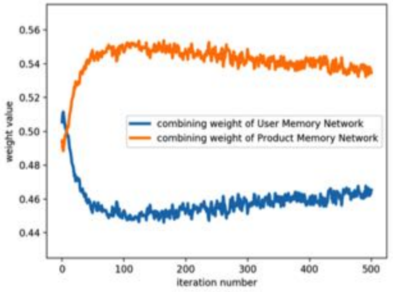




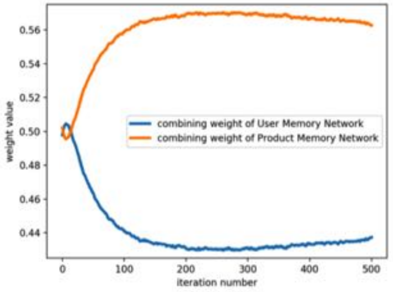
(a) for IMDB dataset

IMDB		Yelp13		Yelp14	
w'_U	w'_P	w'_U	w'_P	w'_U	w'_P
0.534	0.466	0.475	**0.525**	0.436	**0.564**

Average joint weight for three datasets



(b) for Yelp13 dataset



(c) for Yelp14 dataset

Joint weights for three datasets

- Verified the hypothesis

Importance of User vs Product Memory Network

Investigating by Word Frequency Plotting

For IMDB dataset

(a) 10 users who give average highest ratings

(b) 10 users who give average lowest ratings

(a) 10 movies with average highest ratings

(b) 10 movies with average lowest ratings

10 **users** give average **highest/lowest** rating score

10 **movies** have average **highest/lowest** rating score

Importance of User vs Product Memory Network

Investigating by Word Frequency Plotting

For IMDB dataset

(a) 10 users who give average highest ratings

(b) 10 users who give average lowest ratings

(a) 10 movies with average highest ratings

(b) 10 movies with average lowest ratings

For movies reviews

- Users' words are very different
- Products' words are very objective

Importance of User vs Product Memory Network

Investigating by Word Frequency Plotting For Yelp dataset

(a) 10 users who give average highest ratings

(b) 10 users who give average lowest ratings

(a) 10 restaurants with average highest ratings

(b) 10 restaurants with average lowest ratings

For restaurants reviews

- Users' words are not distinguishable
- Products' words shows the sentiments

Number of Hops (Computational Layers)

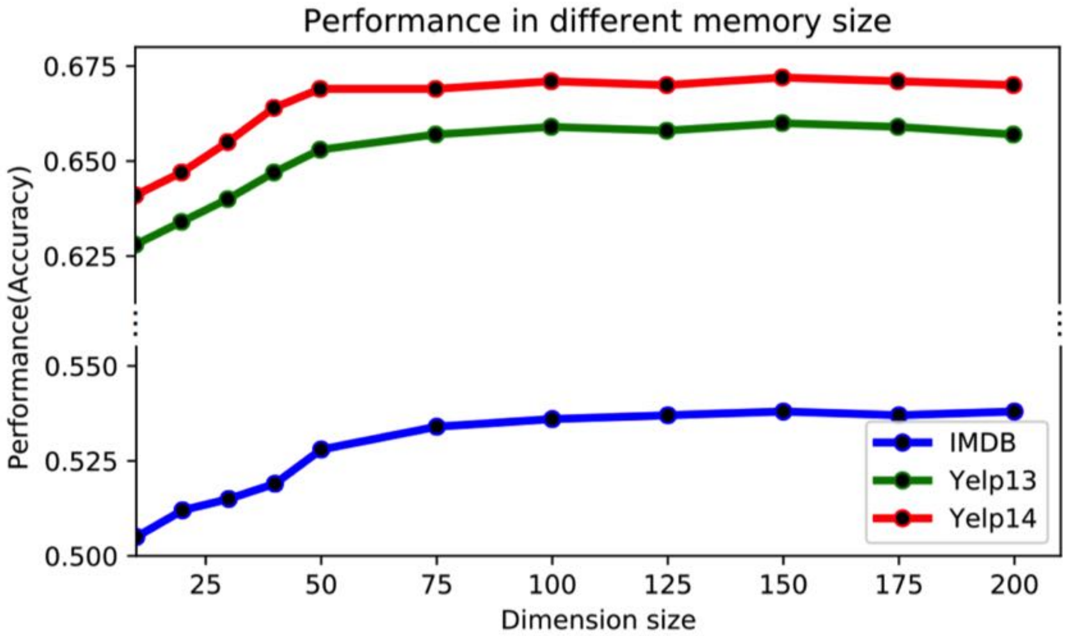
Experimental Results

	IMDB			Yelp13			Yelp14		
	Acc	RMSE	MAE	Acc	RMSE	MAE	Acc	RMSE	MAE
JUPMN-U(1)	0.536	**1.283**	0.737	0.656	0.687	0.380	0.667	0.655	0.361
JUPMN-U(2)	0.526	1.285	0.748	0.653	0.689	0.382	0.665	0.661	0.369
JUPMN-U(3)	0.524	1.295	0.754	0.651	0.692	0.388	0.661	0.667	0.374
JUPMN-P(1)	0.523	1.346	0.769	0.660	0.668	**0.370**	0.670	0.649	0.357
JUPMN-P(2)	0.517	1.348	0.775	0.656	0.680	0.380	0.667	0.656	0.364
JUPMN-P(3)	0.512	1.356	0.661	0.651	0.699	0.388	0.661	0.661	0.370
JUPMN(1)	**0.539**	**1.283**	**0.725**	**0.662**	**0.667**	0.375	**0.676**	**0.641**	**0.351**
JUPMN(2)	0.522	1.299	0.758	0.650	0.700	0.390	0.667	0.650	0.359
JUPMN(3)	0.502	1.431	0.830	0.653	0.686	0.382	0.658	0.668	0.371

Observations

- **Smaller** hop works **better**
- Possible explanations
    - Data distortion
    - Over-fitting

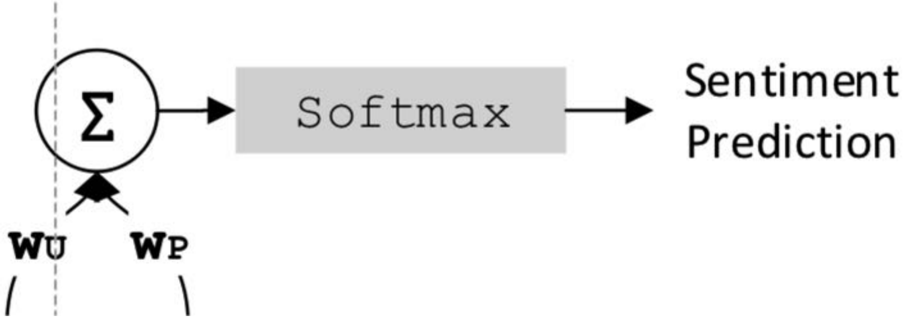
Memory Size



Memory Size	IMDB Acc	IMDB RMSE	IMDB MAE	Yelp13 Acc	Yelp13 RMSE	Yelp13 MAE	Yelp14 Acc	Yelp14 RMSE	Yelp14 MAE
10	0.501	1.572	0.892	0.625	0.788	0.467	0.647	0.692	0.397
20	0.503	1.550	0.866	0.631	0.778	0.456	0.651	0.684	0.384
30	0.516	1.383	0.791	0.643	0.707	0.397	0.668	0.661	0.362
40	0.524	1.367	0.778	0.647	0.695	0.390	0.674	0.641	0.351
50	0.528	1.368	0.769	0.654	0.680	0.379	0.671	0.653	0.356
75	0.529	1.339	0.768	0.655	0.690	0.384	0.674	0.653	0.354
100	0.539	1.283	0.725	0.662	0.667	0.375	0.676	0.641	0.351

- **Larger** memory helps
- When memory size reaches 75, no longer improve
  - There is not enough documents

Joint Weights



JUPMN (not weighted)

$$Output_{JUPMN} = \vec{W}_U \vec{d}_K^u + \vec{W}_P \vec{d}_K^p$$

JUPMN

$$Output_{JUPMN} = w_U \vec{W}_U \vec{d}_K^u + w_P \vec{W}_P \vec{d}_K^p$$

	IMDB			Yelp13			Yelp14		
Model	Acc	RMSE	MAE	Acc	RMSE	MAE	Acc	RMSE	MAE
JUPMN(not weighted)	0.538	1.289	0.737	0.656	0.682	0.379	0.670	0.645	0.354
JUPMN	**0.539**	**1.283**	**0.725**	**0.662**	**0.667**	**0.375**	**0.676**	**0.641**	**0.351**

- Weighted version works better
- Weight help to balance the influences of UMN and PMN

Case Study

Example: Example document

True sentiment label: 10 (most positive)

Predicted sentiment by LSTM network: 1 (most negative)

Predicted sentiment by JUPMN: 10 (most positive)

Original review text:

okay, there are two types of movie lovers: ... they expect to see a Titanic every time they go to the cinema ... this movie sucks? ... it is definitely better than other sci-fi films ..... the audio and visual effects are simply terrific and Travolta's performance is brilliant-funny and interesting. what people expect from sci-fi movies is beyond me ... the rating for Battlefield Earth is below 2.5, which is unacceptable for a movie with such craftsmanship. Scary movie, possibly the worst movie of all time - including home made movies, has a 6! maybe we should all be a little more subtle when we criticize movies like this and especially sci-fi movies, since they have become an endangered genre ... give this movie the recognition it deserves.

- What is this user's opinion?
  - Cite negative reviews to praise
- JUPMN can learn the features of this **user**
  - This user is a science fiction movie
- JUPMN can learn the features of this **movie**(product)
  - This movie is relative great according to other reviews

Conclusion and Future Work

Conclusion

- Proposed JUPMN
- JUPMN outperforms the state-of-the-art sentiment analysis model
- Analysis on different configuration is employed
- Research paper

  Yunfei Long*, Mingyu Ma*, Rong Xiang, Qin Lu, Chu-Ren Huang. **Fusing User Memory and Product Memory for Sentiment Classification.** (*: Equal contribution)

Future Work

- More knowledge in memory network
- Application of JUPMN in more languages datasets

References

Sanjeev Arora, Yingyu Liang, and Tengyu Ma. A simple but tough-to-beat baseline for sentence embeddings. 2016.

Huimin Chen, Maosong Sun, Cunchao Tu, Yankai Lin, and Zhiyuan Liu. Neural sentiment classification with user and product attention. In *EMNLP*, pages 1650–1659, 2016.

Qiming Diao, Minghui Qiu, Chao-Yuan Wu, Alexander J Smola, Jing Jiang, and Chong Wang. Jointly modeling aspects, ratings and sentiments for movie recommendation (jmars). In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 193–202. ACM, 2014.

Zi-Yi Dou. Capturing user and product information for document level sentiment analysis with deep memory network. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 532–537, 2017.

Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, and Chih-Jen Lin. Liblinear: A library for large linear classification. *Journal of machine learning research*, 9(Aug):1871–1874, 2008.

Ronen Feldman. Techniques and applications for sentiment analysis. *Communications of the ACM*, 56(4):82–89, 2013.

Wenliang Gao, Naoki Yoshinaga, Nobuhiro Kaji, and Masaru Kitsuregawa. Modeling user leniency and product popularity for sentiment classification. In *IJCNLP*, pages 1107–1111, 2013.

Palash Goyal and Emilio Ferrara. Graph embedding techniques, applications, and performance: A survey. *arXiv preprint arXiv:1705.02801*, 2017.

Lin Gui, Ruifeng Xu, Yulan He, Qin Lu, and Zhongyu Wei. Intersubjectivity and sentiment: From language to knowledge. In *IJCAI*, pages 2789–2795, 2016.

Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

Yoon Kim. Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*, 2014.

Quoc Le and Tomas Mikolov. Distributed representations of sentences and documents. In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, pages 1188–1196, 2014.

Cheng Li, Xiaoxiao Guo, and Qiaozhu Mei. Deep memory networks for attitude identification. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pages 671–680. ACM, 2017.

Yunfei Long, Lu Qin, Rong Xiang, Minglei Li, and Chu-Ren Huang. A cognition based attention model for sentiment analysis. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 473–482, 2017.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013a.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119, 2013b.

Abhijit Mishra, Diptesh Kanojia, Seema Nagar, Kuntal Dey, and Pushpak Bhattacharyya. Leveraging cognitive features for sentiment analysis. *arXiv preprint arXiv:1701.05581*, 2017.

References

Hamid Palangi, Li Deng, Yelong Shen, Jianfeng Gao, Xiaodong He, Jianshu Chen, Xinying Song, and Rabab Ward. Deep sentence embedding using long short-term memory networks: Analysis and application to information retrieval. *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, 24(4):694–707, 2016.

Bo Pang, Lillian Lee, et al. Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2):1–135, 2008.

Richard Socher, Alex Perelygin, Jean Wu, Jason Chuang, Christopher D Manning, Andrew Ng, and Christopher Potts. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 conference on empirical methods in natural language processing*, pages 1631–1642, 2013.

Sainbayar Sukhbaatar, Jason Weston, Rob Fergus, et al. End-to-end memory networks. In *Advances in neural information processing systems*, pages 2440–2448, 2015.

Zhida Sun, Mingfei Sun, Nan Cao, and Xiaojuan Ma. Videoforest: interactive visual summarization of video streams based on danmu data. In *SIGGRAPH ASIA 2016 Symposium on Visualization*, page 10. ACM, 2016.

Duyu Tang, Furu Wei, Nan Yang, Ming Zhou, Ting Liu, and Bing Qin. Learning sentiment-specific word embedding for twitter sentiment classification. In *ACL (1)*, pages 1555–1565, 2014.

Duyu Tang, Bing Qin, and Ting Liu. Document modeling with gated recurrent neural network for sentiment classification. In *Proceedings of the 2015 conference on empirical methods in natural language processing*, pages 1422–1432, 2015a.

Duyu Tang, Bing Qin, and Ting Liu. Learning semantic representations of users and products for document level sentiment classification. In *ACL (1)*, pages 1014–1023, 2015b.

Duyu Tang, Bing Qin, and Ting Liu. Aspect level sentiment classification with deep memory network. *arXiv preprint arXiv:1605.08900*, 2016.

Jian Tang, Meng Qu, Mingzhe Wang, Ming Zhang, Jun Yan, and Qiaozhu Mei. Line: Large-scale information network embedding. In *Proceedings of the 24th International Conference on World Wide Web*, pages 1067–1077. International World Wide Web Conferences Steering Committee, 2015c.

Cunchao Tu, Han Liu, Zhiyuan Liu, and Maosong Sun. Cane: Context-aware network embedding for relation modeling. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, pages 1722–1731, 2017.

Sida Wang and Christopher D Manning. Baselines and bigrams: Simple, good sentiment and topic classification. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2*, pages 90–94. Association for Computational Linguistics, 2012.

Jason Weston, Sumit Chopra, and Antoine Bordes. Memory networks. *arXiv preprint arXiv:1410.3916*, 2014.

Jiacheng Xu, Danlu Chen, Xipeng Qiu, and Xuangjing Huang. Cached long short-term memory neural networks for document-level sentiment classification. *arXiv preprint arXiv:1610.04989*, 2016.

Ali Yadollahi, Ameneh Gholipour Shahraki, and Osmar R Zaiane. Current state of text sentiment analysis from opinion to emotion mining. *ACM Computing Surveys (CSUR)*, 50(2):25, 2017.

Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alexander J Smola, and Eduard H Hovy. Hierarchical attention networks for document classification. In *HLT-NAACL*, pages 1480–1489, 2016.

Thanks!

A Sentiment Analysis Method To Better Utilize User Profile and Product Information

Mingyu MA (Derek) supervised by Prof. Qin LU

The Hong Kong Polytechnic University
香港理工大學

Department of Computing
電子計算學系